

# On adaptive discretization schemes for the regularization of ill-posed problems with semiiterative methods

W. ERB<sup>(†)</sup>

E.V. SEMENOVA<sup>(\*)</sup>

(†) *Institute of Mathematics, University of Lübeck  
Ratzeburger Allee 160, 23562 Lübeck, Germany*

(\*) *Institute of Mathematics, National Academy of Sciences of Ukraine  
Tereshchenkivska Str. 3, 01601 Kiev, Ukraine*

## Abstract

In this paper we investigate an adaptive discretization strategy for ill-posed linear problems combined with a regularization from a class of semiiterative methods. We show that such a discretization approach in combination with a stopping criterion as the discrepancy principle or the balancing principle yields an order optimal regularization scheme and allows to reduce the computational costs.

**Key words:** linear ill-posed problem, semiiterative method, discrepancy principle, balancing principle, Hölder-type source condition, adaptive discretization scheme

## 1 Introduction

Cost-efficient discretization methods for linear ill-posed equations are well-studied at the present day. Already in the early monographs of A. N. Tikhonov, V. K. Ivanov and M. M. Lavrentiev (see for example [20]) devoted to the regularization of unstable equations different discretizations were studied by applying finite-dimensional approximation techniques to the input data. But only in the beginning of the nineties R. Plato and G. M. Vainikko [15] established estimates on the minimal rank of finite-dimensional operators that guarantee an efficient discretization and preserve the accuracy of the method at the same time. After this work the question on how to reduce the volume of necessary discrete information was studied in several works. A first answer was given in [11]. In this work, as a discretization domain the author proposed the so-called hyperbolic cross for a particular class of ill-posed problems. This cost-efficient discretization approach was then extended to larger classes of ill-posed problems in [13], [17] and several other works. In particular, it is shown in [18] that if the solution of the ill-posed linear equation is smooth enough a large class of regularization methods is more economic when using the hyperbolic cross discretization instead of the classical discretization on a whole rectangular domain.

However, the results of these articles also imply that there are some cases when the discretization with the hyperbolic cross doesn't give an advantage in comparison with the classical discretization. In the case that the smoothness of the solution is low the volume of necessary discrete information is the same for both mentioned approaches. Nevertheless, it is shown in [8] that an additional adaptive strategy for the hyperbolic cross discretization allows to reduce the volume of discrete information even for a low smoothness of the solution. In [8] such a technique was applied for the Tikhonov-Philips regularization. For a particular semiiterative method, namely for the 1-method, a similar adaptive discretization approach was considered in [19] by S. G. Solodky and E. A. Volynets.

The work presented in this article is an extension of the results in [8] and [19]. We show that with a similar adaptive discretization strategy as in [19] it is possible to obtain cost-efficient and

order optimal regularization schemes for a general class of semiiterative methods. Compared to [19], slight changes in the adaptive algorithm allow us to control the influence of the discretization in the approximation error and to get explicit constants in the error estimates. Moreover, beside the discrepancy principle we consider in this work also the balancing principle as a stopping rule for the adaptive algorithm.

In the following two sections we introduce all preliminary information regarding semiiterative regularization methods and projection methods for the hyperbolic cross discretization of linear ill-posed equations. The main results of this article can then be found in Section 4. Here, we present two adaptive regularization schemes (Algorithm 1 and Algorithm 2) for the solution of ill-posed linear equations that combine regularization with semiiterative methods and a cost-efficient discretization strategy based on the hyperbolic cross. The algorithms are stopped either with the discrepancy principle of Morozov or with the balancing principle. In Theorem 4.1 and Theorem 4.3 we will show that both algorithms are order optimal regularization methods if the smoothness of the solution is contained in a given range. In the last section some numerical tests are provided that confirm the theoretical results on the order optimality of the algorithms.

## 1.1 Preliminaries

We shortly introduce the theoretical setup of this article. The space  $X$  denotes a Hilbert space with inner product  $(\cdot, \cdot)$  and norm  $\|x\| = \sqrt{(x, x)}$ . We consider operator equations of the first kind, i.e.

$$Ax = f, \quad (1.1)$$

where  $A : X \rightarrow X$  is a linear compact operator with  $\text{Range}(A) \neq \overline{\text{Range}(A)}$  and  $f \in \text{Range}(A)$ .

By  $x^\dagger$  we denote the solution of (1.1) with minimal norm in  $X$  that satisfies a Hölder-type source condition, namely we suppose that

$$x^\dagger \in M_{\mu, \rho}(A) = \{x \in X : x = |A|^\mu v, \|v\| \leq \rho\}. \quad (1.2)$$

In this article,  $\rho > 0$  is supposed to be known whereas the unknown smoothness parameter  $\mu$  is supposed to belong to an interval  $(0, \mu_0]$  with some  $0 < \mu_0 < \infty$ . Further, we set  $|A| = (A^*A)^{1/2}$  with  $A^*$  denoting the adjoint operator of  $A$ . Also, we suppose that instead of the right-hand side  $f$  in (1.1) we have given some perturbation  $f_\delta$ ,  $\|f - f_\delta\| \leq \delta$  with known noise level  $\delta$ .

Finally, we introduce the class of operators under consideration. We denote by  $\mathcal{H}^r$  the class of compact linear operators  $A$ ,  $\|A\| \leq 1$ , such that for any  $m \in \mathbb{N}$  the following conditions hold:

$$\|(I - P_m)A\| \leq m^{-r}, \quad \|A(I - P_m)\| \leq m^{-r},$$

where  $P_m$  is the orthoprojector on the linear span of the first  $m$  elements of some orthonormal basis  $E = \{e_k\}_{k=1}^\infty$  in the space  $X$ .

## 2 Properties of semiiterative methods

In order to obtain cost-effective iterative schemes for the solution of (1.1), it is useful to consider sequences of orthogonal polynomials and use its three term recurrence relation to generate the iterates (see [6]). Let therefore  $P_k(x)$ ,  $k \in \mathbb{N}_0$ , denote monic polynomials of degree  $k$  that are orthogonal with respect to a weight function supported on the interval  $[-1, 1]$ . Then, the following recursion formula is valid (cf. [2, I. Theorem 4.1]):

$$P_{k+1}(x) = (x - \alpha_k)P_k(x) - \beta_k P_{k-1}(x), \quad P_0(x) = 1, \quad P_1(x) = x - \alpha_0. \quad (2.1)$$

The recursion coefficients  $\alpha_k \in \mathbb{R}$  and  $\beta_k > 0$  are uniquely determined and  $P_k(1) > 0$  holds for all  $k \in \mathbb{N}_0$ . Then (see [5, Algorithm 1] and [6, Theorem 2.1]) a semiiterative method can be defined as

$$\begin{aligned} x_k &= x_{k-1} + ((1 - \alpha_k)\omega_k - 1)(x_{k-1} - x_{k-2}) + 2\omega_k A^*(f - Ax_{k-1}), \quad k \geq 1, \\ x_{-1} &= 0, \quad x_0 = 2\omega_0 A^*f, \end{aligned} \quad (2.2)$$

with

$$\omega_k = \frac{1}{1 - \alpha_k - \beta_k \omega_{k-1}} \quad \text{for } k \geq 1 \quad \text{and} \quad \omega_0 = \frac{1}{1 - \alpha_0}.$$

The iterative method (2.2) yields an approximate solution of (1.1) that can be written as

$$x_k = g_k(A^*A)A^*f, \quad (2.3)$$

where  $g_k(\lambda)$  is a polynomial of degree  $k-1$ . Combining (2.3) and (1.1), we can write the difference  $x^\dagger - x_k$  as

$$x^\dagger - x_k = r_k(A^*A)x^\dagger,$$

with the residual polynomial  $r_k(\lambda) = 1 - \lambda g_k(\lambda)$  of degree  $k$ . For the scheme (2.2) the residual polynomials have the form

$$r_k(\lambda) = \frac{P_k(1 - 2\lambda)}{P_k(1)}, \quad (2.4)$$

with the monic polynomials  $P_k(\lambda)$  satisfying the recursion formula (2.1).

Taking into account that we are only given a perturbed right-hand side  $f_\delta$ , the semiiterative method (2.2) yields

$$x_k^\delta = g_k(A^*A)A^*f_\delta \quad (2.5)$$

as approximate solution of equation (1.1).

For an optimal speed of convergence to the solution  $x^\dagger \in M_{\mu,\rho}$ ,  $0 < \mu \leq \mu_0 < \infty$ , a sufficient condition for the residual polynomials  $r_k$  in (2.4) is (see [4, Section 6.2.] )

$$\begin{aligned} (i) \quad & |r_k(\lambda)| \leq \kappa_0 \quad \text{for all } \lambda \in [0, 1], \quad k \in \mathbb{N}, \\ (ii) \quad & |\lambda^{\frac{\mu}{2}} r_k(\lambda)| \leq \frac{\kappa_\mu}{(k+1)^\mu} \quad \text{for all } 0 < \mu \leq \mu_0, \quad \lambda \in [0, 1], \quad k \in \mathbb{N}, \end{aligned} \quad (2.6)$$

with positive constants  $\kappa_\mu > 0$ ,  $0 < \mu \leq \mu_0$ , and  $\kappa_0 \geq 1$ . The number  $\mu_0$  is called the qualification of the semiiterative method (2.2).

**Example 1:** The  $\nu$ -methods of Brakhage [1] are semiiterative methods based on the monic Jacobi polynomials  $P_k^{(2\nu-\frac{1}{2}, -\frac{1}{2})}$ . For  $\nu > 0$ , its residual polynomials are given as

$$r_k(\lambda) = \frac{P_k^{(2\nu-\frac{1}{2}, -\frac{1}{2})}(1 - 2\lambda)}{P_k^{(2\nu-\frac{1}{2}, -\frac{1}{2})}(1)}.$$

The qualification of this method is  $\mu_0 = 2\nu$ . It is well-known that the Jacobi polynomials satisfy (see [6, Section 4])

$$|r_k(\lambda)| \leq |r_k(0)| = 1, \quad \text{and} \quad |\lambda^\nu r_k(\lambda)| \leq |r_k(1)| = \binom{k+2\nu}{k}^{-1}.$$

Therefore  $\kappa_0 = 1$ , and, if we assume that  $2\nu$  is an integer, we get

$$|\lambda^\nu r_k(\lambda)| \leq \frac{(2\nu)!k!}{(k+2\nu)!} \leq \frac{(2\nu)!}{(k+1)^{2\nu}}.$$

Hence, in this case condition (ii) in (2.6) is satisfied with  $\kappa_{2\nu} = (2\nu)!$  and  $\kappa_\mu \leq \max(\kappa_0, \kappa_{2\nu}) = (2\nu)!$  for all  $0 < \mu \leq 2\nu$ . The last statement follows from [6, Theorem 4.2].

**Example 2:** The  $\nu$ -method with parameter  $\nu = \frac{1}{2}$  is called Chebyshev method of Nemirovskii and Polyak [10]. It is based on the Chebyshev polynomials  $P_k^{(\frac{1}{2}, -\frac{1}{2})}$  of fourth kind and its qualification is  $\mu_0 = 1$ . For the constants  $\kappa_\mu$  in (2.6) we have  $\kappa_\mu = 1$  for all  $0 \leq \mu \leq 1$ .

**Example 3:** The  $\nu$ -method with parameter  $\nu = 1$  was studied in [19] in combination with cost-efficient adaptive discretization schemes. This scheme has qualification  $\mu_0 = 2$  and the constants  $\kappa_\mu$  in (2.6) can be chosen as  $\kappa_0 = 1$  and  $\kappa_\mu = 2$  for all  $0 < \mu \leq 2$ .

**Example 4:** Modified  $\nu$ -methods based on co-dilated Jacobi polynomials with an additional dilation parameter are investigated in [5]. These modified schemes have the same qualification  $\mu_0 = 2\nu$  as the  $\nu$ -methods. However, for these schemes the constant  $\kappa_0$  is in general larger than 1.

For our further analysis we need the Markov inequality in the form

$$|p'_k(\lambda)| \leq 2\kappa k^2, \quad (2.7)$$

where  $p_k(\lambda)$ ,  $\lambda \in [0, 1]$ , is a polynomial of degree  $k$  and  $\kappa = \max_{\lambda \in [0, 1]} |p_k(\lambda)|$ . Using the Markov inequality (2.7) and the conditions (i) and (ii) in (2.6) it is easy to obtain the following inequalities.

**Lemma 2.1.** *If the polynomials  $r_k(\lambda)$ ,  $k \in \mathbb{N}$ , satisfy (i), then*

$$\sup_{\lambda \in [0, 1]} |g_k(\lambda)| \leq 2\kappa_0 k^2, \quad (2.8)$$

$$\sup_{\lambda \in [0, 1]} |\sqrt{\lambda} g_k(\lambda)| \leq 2\kappa_0 k, \quad (2.9)$$

$$|r_k(\lambda) - r_k(\tau)| \leq 2\kappa_0 k^2 |\lambda - \tau| \quad \text{for all } \lambda, \tau \in [0, 1]. \quad (2.10)$$

Moreover, if condition (ii) is satisfied with  $\mu_0 \geq 2$ , then

$$|\lambda r_k(\lambda) - \tau r_k(\tau)| \leq 2\kappa_2 |\lambda - \tau| \quad \text{for all } \lambda, \tau \in [0, 1]. \quad (2.11)$$

**Proof.** Using the Markov inequality (2.7), condition (i) and the mean value theorem, we immediately get for  $\lambda \in [0, 1]$  the inequality

$$|g_k(\lambda)| = \frac{1 - r_k(\lambda)}{\lambda} \leq \sup_{\lambda \in [0, 1]} |r'_k(\lambda)| \leq 2\kappa_0 k^2.$$

Moreover, we get the inequality

$$\lambda g_k(\lambda)^2 = |\lambda g_k(\lambda)| |g_k(\lambda)| = |1 - r_k(\lambda)| |g_k(\lambda)| \leq (1 + \kappa_0) 2\kappa_0 k^2 \leq 4\kappa_0^2 k^2.$$

These two inequalities imply directly (2.8) and (2.9). Using the same ingredients, we also get (2.10):

$$\frac{r_k(\lambda) - r_k(\tau)}{\lambda - \tau} \leq \sup_{\lambda \in [0, 1]} |r'_k(\lambda)| \leq 2\kappa_0 k^2, \quad \lambda \neq \tau.$$

Finally we show (2.11) in the case  $\lambda \neq \tau$  (the case  $\lambda = \tau$  is evident). Using again the mean value theorem and the Markov inequality (2.7), we have

$$\left| \frac{\lambda r_k(\lambda) - \tau r_k(\tau)}{\lambda - \tau} \right| \leq \max_{\lambda \in [0,1]} |(\lambda r_k(\lambda))'| \leq 2(k+1)^2 \max_{\lambda \in [0,1]} |\lambda r_k(\lambda)|.$$

Since the qualification  $\mu_0$  of the residual polynomial is larger than 2, we obtain due to property (ii) in (2.6) the inequality

$$\left| \frac{\lambda r_k(\lambda) - \tau r_k(\tau)}{\lambda - \tau} \right| \leq 2\kappa_2.$$

□

**Lemma 2.2.** *If the properties (i) and (ii) are satisfied with  $\mu_0 \geq 2$ , then the inequalities*

$$|\lambda(r_k(\lambda) - r_k(\tau))| \leq (\kappa_0 + 2\kappa_2)|\lambda - \tau|, \quad (2.12)$$

and

$$\sqrt{\lambda}|r_k(\lambda) - r_k(\tau)| \leq 2\kappa_0 \sqrt{\frac{1}{2} + \frac{\kappa_2}{\kappa_0}} k |\lambda - \tau| \quad (2.13)$$

hold for  $\lambda, \tau \in [0, 1]$ .

**Proof.** By (2.11) and (i), we immediately obtain

$$|\lambda(r_k(\lambda) - r_k(\tau))| \leq |\lambda r_k(\lambda) - \tau r_k(\tau)| + |\lambda - \tau| |r_k(\tau)| \leq (\kappa_0 + 2\kappa_2)|\lambda - \tau|.$$

The inequality (2.13) follows from (2.12) and (2.10) since

$$\lambda |r_k(\lambda) - r_k(\tau)|^2 = \lambda |r_k(\lambda) - r_k(\tau)| |r_k(\lambda) - r_k(\tau)| \leq 2\kappa_0(\kappa_0 + 2\kappa_2)k^2 |\lambda - \tau|^2.$$

□

We remark that the qualitative statements of Lemma 2.1 and 2.2 are not new. In a simplified version, they can for instance be found in [19] for the residual polynomials of the 1-method. However, for the error estimates in the following sections the explicit constants on the right hand side of the above inequalities, in particular of (2.13), play an important role. Therefore, we decided to include both lemmas with proof in this article.

### 3 Discretization schemes for linear equations

To obtain finite dimensional approximations  $A_\Omega$  of the operator  $A$  we consider in this article projection methods and use the inner products

$$(Ae_i, e_j), (f_\delta, e_j), \quad (i, j) \in \Omega \quad (3.1)$$

as discrete information about the linear equation (1.1), where  $\Omega \subset \{(i, j) \in \mathbb{N}^2\}$  denotes some subdomain of the coordinate plane.

In the following, we denote by  $R_\Omega^{\mu_0}$  the class of iterative methods that solve (1.1) approximatively using the discretization domain  $\Omega$  and a chosen semiiterative method with qualification  $\mu_0 > 0$ . We denote the corresponding iterates by

$$x_{\Omega, k}^\delta := g_k(A_\Omega^* A_\Omega) A_\Omega^* f_\delta, \quad k \in \mathbb{N}. \quad (3.2)$$

We have the following general error bound for the iterates  $x_{\Omega, k}^\delta$ .

**Lemma 3.3.** *Let the solution  $x^\dagger \in M_{\mu,\rho}(A)$  of (1.1) satisfy a Hölder-type source condition with smoothness  $0 < \mu \leq \mu_0$ . Then, for the accuracy of an iterative method in the class  $R_\Omega^{\mu_0}$  we obtain the following estimate:*

$$\|x^\dagger - x_{\Omega,k}^\delta\| \leq \kappa_\mu \rho k^{-\mu} + 2\kappa_0 k \delta + 2\kappa_0 k^2 \|x^\dagger\| \left( \|A^*A - A_\Omega^*A_\Omega\| + \|A_\Omega^*(A_\Omega - A)\| \right).$$

**Proof.** We decompose the total error in the three terms

$$\|x^\dagger - x_{\Omega,k}^\delta\| \leq \|x^\dagger - x_k\| + \|g_k(A_\Omega A_\Omega)A_\Omega^*(f - f_\delta)\| + \|x_k - g_k(A_\Omega^*A_\Omega)A_\Omega^*f\|$$

and estimate each summand separately. For the estimate of the first summand we use the fact that  $x^\dagger \in M_{\mu,\rho}(A)$  and adopt a well-known result for regularization filters (see [16, Lemma 3.3.6]) with qualification  $\mu_0 \geq \mu$ . This gives

$$\|x^\dagger - x_k\| = \kappa_\mu (k+1)^{-\mu} \rho \leq \kappa_\mu \rho k^{-\mu}.$$

The second estimate can be easily deduced from (2.9):

$$\|g_k(A_\Omega A_\Omega)A_\Omega^*(f - f_\delta)\| \leq \sup_{\lambda \in [0,1]} |\lambda^{1/2} g_k(\lambda)| \|f - f_\delta\| \leq 2\kappa_0 k \delta.$$

Finally, we estimate the third summand using (2.8) and (2.10):

$$\begin{aligned} \|x_k - g_k(A_\Omega^*A_\Omega)A_\Omega f\| &\leq \|(g_k(A^*A)A^*A - g_k(A_\Omega^*A_\Omega)A_\Omega^*A_\Omega)x^\dagger\| + \|g_k(A_\Omega^*A_\Omega)A_\Omega^*(A_\Omega - A)x^\dagger\| \\ &\leq \|(r_k(A_\Omega^*A_\Omega) - r_k(A^*A))x^\dagger\| + \|g_k(A_\Omega^*A_\Omega)\| \|A_\Omega^*(A_\Omega - A)x^\dagger\| \\ &\leq 2\kappa_0 k^2 \|x^\dagger\| \|A_\Omega^*A_\Omega - A^*A\| + 2\kappa_0 k^2 \|x^\dagger\| \|A_\Omega^*(A_\Omega - A)\|. \end{aligned}$$

Combining these three estimates, we get the statement of the lemma.  $\square$

We remark that the qualitative statement of Lemma 3.3 can be already found in [19, Lemma 3.3] (therein proven for the 1-method). Since the constants in the error estimate are important for our purposes, we decided to formulate also the above proof explicitly, although it is quite similar to the proof in [19, Lemma 3.3].

The efficiency of the discretization depends heavily on the structure of  $\Omega$ . One standard choice for  $\Omega$  (suggested in [15]) is to take the rectangular domain  $\Omega = [1, \dots, M] \times [1, \dots, N]$ . In this case the discretized operator  $A_\Omega$  has the form

$$A_\Omega = P_M A P_N.$$

A second more efficient possibility is to take the hyperbolic cross

$$\Gamma_n := \bigcup_{k=1}^{2n} (2^{k-1}, 2^k] \times [1, 2^{2n-k}] \cup \{1\} \times [1, 2^{2n}] \subset \mathbb{N}^2 \quad (3.3)$$

as a discretization domain. In this case, the discretized operator  $A_n := A_{\Gamma_n}$  has the form

$$A_n := A_{\Gamma_n} = \sum_{k=1}^{2n} (P_{2^k} - P_{2^{k-1}}) A P_{2^{2n-k}} + P_1 A P_{2^{2n}}. \quad (3.4)$$

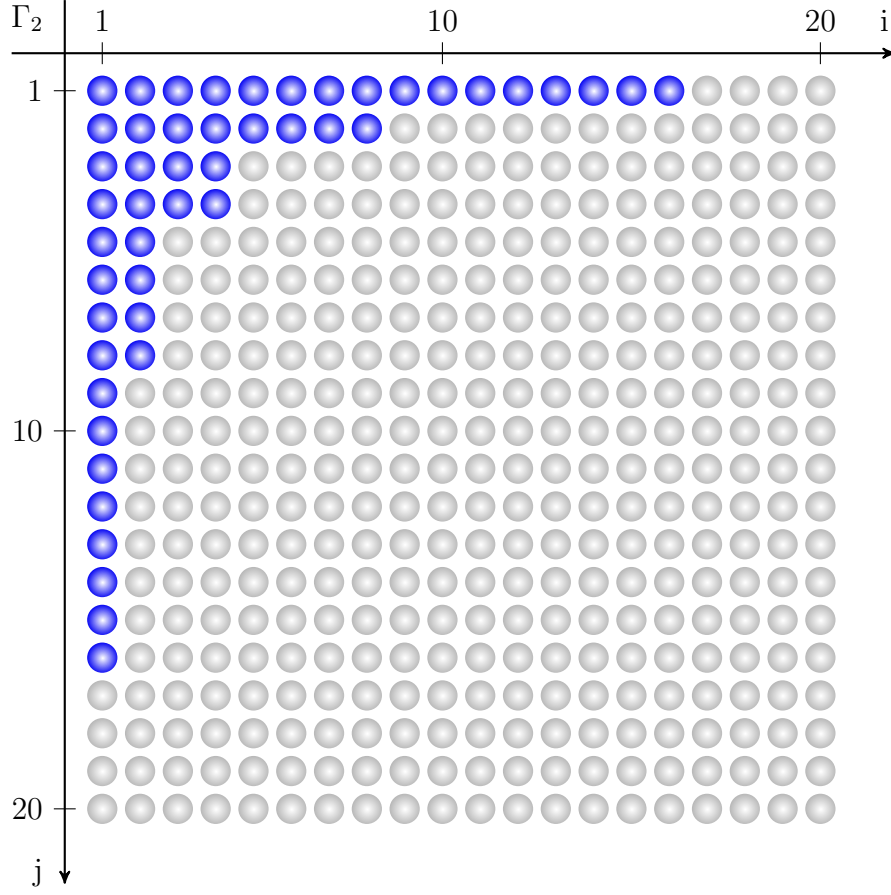


Figure 1: Visualization of the hyperbolic cross  $\Gamma_2$ . The elements contained in  $\Gamma_2$  are colored in blue.  $\Gamma_2$  contains 48 elements. In contrast, the rectangular domain  $\Omega = [1, \dots, 2^4] \times [1, \dots, 2^4]$  contains 256 elements.

**Remark 3.1.** *It is easy to compute the volume of Galerkin information necessary for the realization of the hyperbolic cross approximation  $A_n$  given by (3.4): one has to compute  $\#\Gamma_n = 2^{2n}(n+1)$  inner products to construct  $A_n$ . In comparison, one has to compute  $2^{4n}$  inner products for the approximation with  $A_\Omega$  for the standard quadratic domain  $\Omega = [1, \dots, N]^2$ ,  $N = 2^{2n}$ .*

If  $A \in \mathcal{H}^r$  and  $A_n$  has the form (3.4), the following error estimates hold (see [17]):

$$\|A^*A - A_n^*A_n\| \leq (1 + 2^{r+3})2^{-2rn}n, \quad \|A_n^*(A - A_n)\| \leq 3 \cdot 2^{-2rn+r}n. \quad (3.5)$$

Moreover, since  $A \in \mathcal{H}^r$  we have the bound

$$\|A - AP_{2^{2n}}\| \leq 2^{-2rn}. \quad (3.6)$$

**Corollary 3.1.** *If  $A \in \mathcal{H}^r$ ,  $A_n$  has the form (3.4) and*

$$(1 + 2^{r+3})2^{-2rn}n < \frac{\gamma\delta}{2k\rho} \quad (3.7)$$

with  $k \in \mathbb{N}$  and a control parameter  $\gamma > 0$ , then

$$\|A^*A - A_n^*A_n\| \leq \frac{\gamma\delta}{2k\rho}, \quad \|A_n^*(A - A_n)\| \leq \frac{\gamma\delta}{2k\rho}, \quad \|A - AP_{2^{2n}}\| \leq \frac{\gamma\delta}{2k\rho n}. \quad (3.8)$$

If  $x^\dagger \in M_{\mu,\rho}(A)$ , the approximation error of methods from the class  $R_n^{\mu_0} := R_{\Gamma_n}^{\mu_0}$  with  $\mu_0 \geq \mu > 0$  is given by

$$\|x^\dagger - x_{n,k}^\delta\| \leq \kappa_\mu \rho k^{-\mu} + 2\kappa_0 k \delta (1 + \gamma). \quad (3.9)$$

In the adaptive algorithms of the next section the values  $\delta$ ,  $\rho$  and  $k \in \mathbb{N}$  can not be chosen freely. The control parameter  $\gamma$  in (3.7) enables a possible user to handle the trade-off between accuracy and cost-efficiency of the adaptive scheme.

## 4 Regularization with semiiterative methods and adaptive discretization strategies

This section contains the main new results of the article. We present two algorithms in which the regularization with semiiterative methods is combined with an adaptive and cost-efficient discretization strategy. As stopping rule for the regularization we consider the discrepancy principle of Morozov [9] on the one hand and the balancing principle [12] on the other.

### 4.1 The discrepancy principle as stopping rule

To solve the linear ill-posed problem (1.1) we consider the adaptive Algorithm 1 which combines a method from the class  $R_n^{\mu_0}$ ,  $\mu_0 \geq 2$ , with an adaptive discretization strategy and the discrepancy principle of Morozov [9] (implemented as (4.2) in Algorithm 1) as stopping rule.

In Theorem 4.1 we show that Algorithm 1 yields an order optimal regularization scheme for the solution of (1.1). For the proof we need some additional statements.

**Lemma 4.4.** *Let  $A \in \mathcal{H}^r$ ,  $x^\dagger \in M_{\mu,\rho}(A)$  and  $x_{n,k}^\delta$  be computed according to Algorithm 1. Then, we have the estimate*

$$\|Ax_k - f\| \leq \|A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta\| + c_1 \delta,$$

with  $c_1 = \kappa_0 + 2 + \left(\sqrt{\kappa_0\left(\frac{\kappa_0}{2} + \kappa_2\right)} + \frac{1}{2n}\right) \gamma$ .

**Proof.** To prove the statement we write  $Ax_k - f$  in the following telescoping sum:

$$Ax_k - f = Ax_k - Ax_k^\delta + Ax_k^\delta - A_n x_{n,k}^\delta + A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta + P_{2^{2n}} f_\delta - P_{2^{2n}} f + P_{2^{2n}} f - f$$

Now, using the triangle inequality and the definition of the generating polynomials  $g_n$  we get

$$\begin{aligned} \|Ax_k - f\| &\leq \|Ag_k(A^*A)A^*(f - f_\delta)\| + \|Ag_k(A^*A)A^*f_\delta - A_n g_k(A_n^* A_n)A_n^* f_\delta\| \\ &\quad + \|A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta\| + \|P_{2^{2n}}(f_\delta - f)\| + \|P_{2^{2n}} f - f\|. \end{aligned}$$

Since  $g_k(A^*A)A^* = A^*g_k(AA^*)$ , we can further estimate

$$\begin{aligned} \|Ax_k - f\| &\leq \sup_{\lambda \in [0,1]} |\lambda g_k(\lambda)| \|f - f_\delta\| + \|(r_k(AA^*) - r_k(A_n A_n^*))Ax^\dagger\| \\ &\quad + \|A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta\| + \|f_\delta - f\| + \|(P_{2^{2n}} A - A)x^\dagger\|. \end{aligned}$$

Using property (i) of the residual polynomials, inequality (2.13) as well as the inequalities (3.8) in Corollary 3.1, we finally obtain

$$\begin{aligned} \|Ax_k - f\| &\leq (\kappa_0 + 2)\delta + 2\sqrt{\kappa_0\left(\frac{\kappa_0}{2} + \kappa_2\right)} k \|A^*A - A_n^* A_n\| \rho + \|A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta\| + \|P_{2^{2n}} A - A\| \rho \\ &\leq \|A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta\| + \left(\kappa_0 + 2 + \left(\sqrt{\kappa_0\left(\frac{\kappa_0}{2} + \kappa_2\right)} + \frac{1}{2n}\right) \gamma\right) \delta. \end{aligned}$$

□



---

**Algorithm 1** Adaptive algorithm to solve (1.1) using the discrepancy principle

---

given data  $A \in \mathcal{H}^r, \delta, f_\delta, \rho$ .

choose control parameters  $\gamma > 0, \tau > \kappa_0 \left(1 + \sqrt{\frac{1}{2} + \frac{\kappa_2}{\kappa_0} \gamma}\right)$ .

choose discretization level  $n \geq 1$  such that  $(1 + 2^{r+3})2^{-2rn}n < \frac{\gamma\delta}{2\rho}$  holds.

compute Galerkin information:

$$(f_\delta, e_j), \quad j \in [1, 2^{2n}], \quad (Ae_i, e_j), \quad (i, j) \in \Gamma_n.$$

**while** (discrepancy principle = false) **do**

choose  $K_n \in \mathbb{N}$  as maximal integer such that

$$(1 + 2^{r+3})2^{-2rn}n < \frac{\gamma\delta}{2K_n\rho} \tag{4.1}$$

is satisfied.

**for** ( $k = 1 : K_n$ ) **do**

compute  $k$ -th. iterate of semiiterative method in the class  $R_n^{\mu_0}, \mu_0 \geq 2$  (cf. (2.2)):

$$x_{n,k}^\delta = x_{n,k-1}^\delta + ((1 - \alpha_k)\omega_k - 1)(x_{n,k-1}^\delta - x_{n,k-2}^\delta) + 2\omega_k A_n^*(f_\delta - A_n x_{n,k-1}^\delta).$$

**if**

$$\|A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta\| \leq \tau\delta \quad \text{and} \quad \|A_n x_{n,j}^\delta - P_{2^{2n}} f_\delta\| > \tau\delta \quad \text{for all } j < k, \tag{4.2}$$

**then**

discrepancy principle = true,

**return** stopping index  $K = k$ , discretization level  $n$  and solution  $x_{n,K}^\delta$ .

**end if**

**end for**

increase discretization level  $n \rightarrow n + 1$ .

compute new Galerkin information:

$$(f_\delta, e_j), \quad j \in [2^{2n-2}, 2^{2n}]; \quad (Ae_i, e_j), \quad (i, j) \in \Gamma_n \setminus \Gamma_{n-1}.$$

**end while**

---

**Lemma 4.5.** *Let  $A \in \mathcal{H}^r$ ,  $x^\dagger \in M_{\mu,\rho}(A)$ ,  $0 < \mu \leq \mu_0 - 1$ ,  $\mu_0 \geq 2$  and  $x_{n,k}^\delta$  be computed according to Algorithm 1. Further, we assume that the level of noise satisfies  $\delta < \|f\|$ . Then, the stopping index  $K$  is bounded by*

$$K < c_2 \rho^{\frac{1}{\mu+1}} \delta^{-\frac{1}{\mu+1}}, \quad \text{with} \quad c_2 = \max \left\{ \left( \frac{\kappa_\mu}{\tau - \kappa_0 \left( 1 + \sqrt{\frac{1}{2} + \frac{\kappa_2}{\kappa_0} \gamma} \right)} \right)^{\frac{1}{\mu+1}}, 1 \right\}. \quad (4.3)$$

If Algorithm 1 is not stopped in the first iteration of the while loop, then also the index  $K_{n-1}$  satisfies the estimate

$$K_{n-1} + 1 < c_2 \rho^{\frac{1}{\mu+1}} \delta^{-\frac{1}{\mu+1}}. \quad (4.4)$$

**Proof.** To prove this statement we consider the second inequality (4.2) in Algorithm 1. For the stopping index  $K \geq 2$ , we get

$$\begin{aligned} \tau \delta &< \|A_n x_{n,K-1}^\delta - P_{2^{2n}} f_\delta\| = \|r_{K-1}(A_n A_n^*) f_\delta\| \\ &\leq \|r_{K-1}(A A^*) A x^\dagger\| + \|(r_{K-1}(A_n A_n^*) - r_{K-1}(A A^*)) A x^\dagger\| + \|r_{K-1}(A_n A_n^*)(f - f_\delta)\|. \end{aligned}$$

Now using the fact that  $x^\dagger \in M_{\mu,\rho}(A)$ , property (i) of the residual polynomials  $r_n$  as well as inequality (2.13) (here  $\mu_0 \geq 2$  is necessary), we obtain

$$\tau \delta < \| |A|^{\mu+1} r_{K-1}(A^* A) v \| + 2 \sqrt{\kappa_0 \left( \frac{\kappa_0}{2} + \kappa_2 \right)} (K-1) \|A^* A - A_n^* A_n\| \rho + \kappa_0 \delta.$$

Finally, using property (ii) of (2.6) (here  $\mu + 1 \leq \mu_0$  must be satisfied) and Corollary 3.1 (the conditions of the corollary are satisfied by the construction of Algorithm 1) we conclude

$$\tau \delta < \kappa_{\mu+1} K^{-(\mu+1)} \rho + \left( \kappa_0 + \sqrt{\kappa_0 \left( \frac{\kappa_0}{2} + \kappa_2 \right)} \gamma \right) \delta.$$

Solving this inequality for the index  $K$  implies inequality (4.3).

For the case that Algorithm 1 is stopped at  $K = 1$ , the assumption  $\delta < \|f\|$  implies

$$\delta < \|f\| = \| |A|^{\mu+1} v \| \leq \rho.$$

Thus, we get for  $K = 1$

$$K < \rho^{\frac{1}{\mu+1}} \delta^{-\frac{1}{\mu+1}}.$$

The proof of inequality (4.4) follows the same lines of argumentation as the proof of inequality (4.3) with  $K - 1$  replaced by  $K_{n-1}$  and  $n$  replaced by  $n - 1$ .  $\square$

**Theorem 4.1.** *Let  $A \in \mathcal{H}^r$ ,  $\delta < \|f\|$  and  $\mu_0 \geq 2$  for the qualification of the semiiterative method. Then, Algorithm 1 gives an order optimal regularization method for the solution  $x^\dagger$  of (1.1) in the class  $M_{\mu,\rho}(A)$  for all  $0 < \mu \leq \mu_0 - 1$ . In particular, the approximative solution  $x_{n,K}^\delta$  given by Algorithm 1 satisfies*

$$\|x^\dagger - x_{n,K}^\delta\| \leq C \rho^{\frac{1}{\mu+1}} \delta^{\frac{\mu}{\mu+1}}, \quad (4.5)$$

$$\text{with } C = \left( \kappa_0^{\frac{1}{\mu+1}} (\tau + c_1)^{\frac{\mu}{\mu+1}} + 2\kappa_0(1 + \gamma)c_2 \right).$$

**Proof.** We only have to prove inequality (4.5), i.e. that Algorithm 1 gives an order optimal reconstruction scheme for all  $0 < \mu \leq \mu_0 - 1$ . Then, it follows from a general result of R. Plato [14, Theorem 2.1] that Algorithm 1 is also a regularization method for the solution of (1.1). Using the

estimates of Lemma 3.3 and Corollary 3.1 (excluding the estimate for the approximation error), we get the error bound.

$$\|x^\dagger - x_{n,K}^\delta\| \leq \|x^\dagger - x_K\| + 2\kappa_0 K \delta (1 + \gamma). \quad (4.6)$$

To estimate the approximation error, we use an interpolation inequality (see [16, Satz 2.4.2.] or [7, Satz 2.3.3.]) and obtain

$$\begin{aligned} \|x^\dagger - x_K\| &\leq \|r_K(A^*A)x^\dagger\| = \| |A|^\mu r_K(A^*A)v \| \\ &\leq \| |A|^{\mu+1} r_K(A^*A)v \|_{\frac{\mu}{\mu+1}} \|r_K(A^*A)v\|_{\frac{1}{\mu+1}} \\ &\leq \|Ar_K(A^*A)x^\dagger\|_{\frac{\mu}{\mu+1}} (\kappa_0 \rho)^{\frac{1}{\mu+1}} = \|Ax_K - f\|_{\frac{\mu}{\mu+1}} (\kappa_0 \rho)^{\frac{1}{\mu+1}}. \end{aligned}$$

Now, Lemma 4.4 and the discrepancy principle (4.2) give

$$\begin{aligned} \|x^\dagger - x_K\| &\leq (\|A_n x_{n,k}^\delta - P_{2^{2n}} f_\delta\| + c_1 \delta)^{\frac{\mu}{\mu+1}} (\kappa_0 \rho)^{\frac{1}{\mu+1}} \\ &\leq \kappa_0^{\frac{1}{\mu+1}} (\tau + c_1)^{\frac{\mu}{\mu+1}} \rho^{\frac{1}{\mu+1}} \delta^{\frac{\mu}{\mu+1}}. \end{aligned}$$

For the data error in (4.6) Lemma 4.5 gives the bound

$$2\kappa_0 K \delta (1 + \gamma) \leq 2\kappa_0 (1 + \gamma) c_2 \rho^{\frac{1}{\mu+1}} \delta^{\frac{\mu}{\mu+1}}.$$

In total we can conclude:

$$\|x^\dagger - x_{n,K}^\delta\| \leq \left( \kappa_0^{\frac{1}{\mu+1}} (\tau + c_1)^{\frac{\mu}{\mu+1}} + 2\kappa_0 (1 + \gamma) c_2 \right) \rho^{\frac{1}{\mu+1}} \delta^{\frac{\mu}{\mu+1}}.$$

□

Regarding the computational expenses of Algorithm 1, we get the following result.

**Theorem 4.2.** *Let  $A \in \mathcal{H}^r$ ,  $x^\dagger \in M_{\mu,\rho}(A)$ ,  $0 < \mu \leq \mu_0 - 1$ ,  $\mu_0 \geq 2$  and  $x_{n,k}^\delta$  be computed according to Algorithm 1. Further, we assume that the level of noise satisfies  $\delta < \|f\|$  and that Algorithm 1 is not stopped in the first iteration of the while loop. Then, the discretization level  $n$  of the solution  $x_{n,K}^\delta$  is bounded by.*

$$n < c_4 + c_5 \ln \frac{\rho}{\delta} \quad (4.7)$$

with  $c_4 = \frac{1}{r \ln 2} \ln \frac{c_2}{\gamma} \frac{2^{r+1}(1+2^{r+3})}{2^r-1}$  and  $c_5 = \frac{1}{r \ln 2} \frac{\mu+2}{\mu+1}$ . Further, the number of calculated inner products  $\langle Ae_i, e_j \rangle$  for the domain  $\Gamma_n$  can be estimated as

$$\#\Gamma_n = 2^{2n}(n+1) < c_3 \left( \frac{\rho}{\delta} \right)^{\frac{\mu+2}{r(\mu+1)}} \left( 1 + c_4 + c_5 \ln \frac{\rho}{\delta} \right)^{1+\frac{1}{r}}, \quad (4.8)$$

with  $c_3 = \left( \frac{c_2}{\gamma} 2^{2r+1} (1 + 2^{r+3}) \right)^{\frac{1}{r}}$ ,

**Proof.** Since Algorithm 1 is not stopped in the first iteration we have by definition of the index  $K_{n-1}$  in (4.1):

$$(1 + 2^{r+3}) 2^{-2r(n-1)} (n-1) > \frac{\gamma \delta}{2(K_{n-1} + 1)\rho}.$$

This can be formulated equivalently as

$$\frac{2^{2rn}}{n-1} < 2^{2r+1} (1 + 2^{r+3}) \frac{\rho}{\gamma \delta} (K_{n-1} + 1).$$

By inequality (4.4) in Lemma 4.5 we get now the estimate

$$\frac{2^{2rn}}{n-1} < \frac{c_2}{\gamma} 2^{2r+1} (1 + 2^{r+3}) \left(\frac{\rho}{\delta}\right)^{\frac{\mu+2}{\mu+1}}. \quad (4.9)$$

The Bernoulli inequality  $1 + (n-1)(2^r - 1) \leq 2^{r(n-1)}$  yields the bound  $(n-1) \leq \frac{2^{rn-r}}{2^r-1}$ . Thus, we obtain from (4.9) the inequality

$$2^{rn} < \frac{c_2}{\gamma} \frac{2^{r+1}}{2^r-1} (1 + 2^{r+3}) \left(\frac{\rho}{\delta}\right)^{\frac{\mu+2}{\mu+1}}.$$

Now, taking the logarithm on both sides gives the desired inequality (4.7). Finally, using again (4.9) and (4.7) we obtain

$$\#\Gamma_n = 2^{2n}(n+1) < \left(\frac{2^{2rn}}{n-1}\right)^{\frac{1}{r}} (n+1)^{1+\frac{1}{r}} < c_3 \left(\frac{\rho}{\delta}\right)^{\frac{\mu+2}{r(\mu+1)}} \left(1 + c_4 + c_5 \ln \frac{\rho}{\delta}\right)^{1+\frac{1}{r}},$$

with  $c_3 = \left(\frac{c_2}{\gamma} 2^{2r+1} (1 + 2^{r+3})\right)^{\frac{1}{r}}$ . □

**Remark 4.2.** For a standard nonadaptive Galerkin scheme with quadratic domain  $\Omega = [1, \dots, 2^{2n}]^2$  the computational expenses for a suitable large discretization level  $n$  turn out to be of order  $O((\frac{\rho}{\delta})^{\frac{2}{r}})$ . This asymptotic result can be deduced from the error bounds given in [15] and the fact that  $A \in H_r$ . In comparison, by Theorem 4.2 the computational costs of Algorithm 1 are of order  $O\left((\frac{\rho}{\delta})^{\frac{\mu+2}{r(\mu+1)}} (\ln \frac{\rho}{\delta})^{1+\frac{1}{r}}\right)$ . In this sense, the adaptive algorithm presented in this article is more economic than regularization schemes using the standard Galerkin scheme. The same order of complexity was also shown for the adaptive scheme in [8] using a Tikhonov regularization and in [19] for the regularization with the 1-method (with different proofs).

## 4.2 The balancing principle as stopping rule

In this section, we consider a second adaptive algorithm to solve (1.1) combining again a method from the class  $R_n^{\mu_0}$   $\mu_0 \geq 2$  with an adaptive discretization strategy. However, this time we use the balancing principle (see (4.11) in Algorithm 2) as stopping rule.

For technical purposes we need the index  $K_{opt} := \left\lceil \left(\frac{2(1+\gamma)\delta}{\kappa_\mu \rho}\right)^{-\frac{1}{\mu+1}} \right\rceil$ . It is easy to see that  $\kappa_\mu \rho K_{opt}^{-\mu} \leq 2(1+\gamma)K_{opt}\delta$  is satisfied and that for all  $k \geq K_{opt}$  we have the inequality

$$\kappa_\mu \rho k^{-\mu} \leq 2(1+\gamma)k\delta. \quad (4.12)$$

**Theorem 4.3.** Let  $A \in \mathcal{H}^r$  and  $\mu_0 \geq 2$  for the class  $R_n^{\mu_0}$ . Then, Algorithm 2 gives an order optimal regularization method for the solution  $x^\dagger$  of (1.1) in the class  $M_{\mu,\rho}(A)$  for all  $0 < \mu \leq \mu_0$ . In particular, the approximative solution  $x_{n,K}^\delta$  given by Algorithm 2 satisfies

$$\|x^\dagger - x_{n,K}^\delta\| \leq C \rho^{\frac{1}{\mu+1}} \delta^{\frac{\mu}{\mu+1}}, \quad (4.13)$$

with  $C = 12\kappa_\mu^{\frac{1}{\mu+1}} (2(1+\gamma)\kappa_0)^{\frac{\mu}{\mu+1}}$ .

**Proof.** We check that  $K_{opt} \geq K$ . For all  $k \geq 1$ , Corollary 3.1 implies

$$\|x_{n,k}^\delta - x_{n,K_{opt}}^\delta\| \leq \|x^\dagger - x_{n,K_{opt}}^\delta\| + \|x^\dagger - x_{n,k}^\delta\| \leq \kappa_\mu \rho K_{opt}^{-\mu} + 2(1+\gamma)\kappa_0 K_{opt}\delta + \kappa_\mu \rho k^{-\mu} + 2(1+\gamma)\kappa_0 k\delta.$$

---

**Algorithm 2** Adaptive algorithm to solve (1.1) using the balancing principle

---

given data  $A \in \mathcal{H}^r, \delta, f_\delta, \rho$ .

choose control parameters  $\gamma > 0, K_{sec} \in \mathbb{N}$ .

choose discretization level  $n$  such that  $(1 + 2^{r+3})2^{-2rn}n < \frac{\gamma\delta}{2\rho}$  holds.

compute Galerkin information:

$$(f_\delta, e_j), \quad j \in [1, 2^{2n}]; \quad (Ae_i, e_j), \quad (i, j) \in \Gamma_n.$$

**while** (balancing principle = false) **do**

choose  $K_n \in \mathbb{N}$  as maximal integer such that

$$(1 + 2^{r+3})2^{-2rn}n < \frac{\gamma\delta}{2K_n\rho} \quad (4.10)$$

is satisfied.

**for** ( $k = 1 : K_n + K_{sec}$ ) **do**

compute iterates of semiiterative method in the class  $R_n^{\mu_0}, \mu_0 \geq 2$  (cf. (2.2)):

$$x_{n,k}^\delta = x_{n,k-1}^\delta + ((1 - \alpha_k)\omega_k - 1)(x_{n,k-1}^\delta - x_{n,k-2}^\delta) + 2\omega_k A_n^*(f_\delta - A_n x_{n,k-1}^\delta).$$

**end for**

compute the set

$$D_n^+ = \{k : k \leq K_n, \|x_{n,k}^\delta - x_{n,j}^\delta\| \leq 8(1 + \gamma)\kappa_0 j \delta \text{ for all } j \text{ with } k < j \leq K_n + K_{sec}\}. \quad (4.11)$$

**if**  $D_n^+ = \emptyset$  **then**

increase discretization level  $n \rightarrow n + 1$ .

**else**

balancing principle = true

**return** stopping index  $K = \min\{k : k \in D_n^+\}$ , discretization level  $n$  and solution  $x_{n,K}^\delta$

**end if**

compute new Galerkin information:

$$(f_\delta, e_j), \quad j \in [2^{2n-2}, 2^{2n}]; \quad (Ae_i, e_j), \quad (i, j) \in \Gamma_n \setminus \Gamma_{n-1}.$$

**end while**

---

Now, using (4.12) we get for all  $k \geq K_{opt}$ :

$$\|x_{n,k}^\delta - x_{n,K_{opt}}^\delta\| \leq 4(1+\gamma)\kappa_0 k\delta + 4(1+\gamma)\kappa_0 K_{opt}\delta \leq 8(1+\gamma)\kappa_0 k\delta.$$

Thus,  $K_{opt} \in D_n^+$  if  $K_{opt} \leq K_n$  and  $K_{opt} \geq K$ . Therefore, by the balancing principle (4.11) the total error can be bounded as follows:

$$\begin{aligned} \|x^\dagger - x_{n,K}^\delta\| &\leq \|x^\dagger - x_{n,K_{opt}}^\delta\| + \|x_{n,K_{opt}}^\delta - x_{n,K}^\delta\| \leq \kappa_\mu \rho K_{opt}^{-\mu} + 2(1+\gamma)\kappa_0 K_{opt}\delta + 8(1+\gamma)\kappa_0 K_{opt}\delta \\ &\leq 12(1+\gamma)\kappa_0 K_{opt}\delta = 12(1+\gamma)\kappa_0 \left[ \left( \frac{2(1+\gamma)\kappa_0\delta}{\kappa_\mu \rho} \right)^{-\frac{1}{\mu+1}} \right] \delta \\ &\leq 24(1+\gamma)\kappa_0 \left( \frac{2(1+\gamma)\kappa_0\delta}{\kappa_\mu \rho} \right)^{-\frac{1}{\mu+1}} \delta = C \rho^{\frac{1}{\mu+1}} \delta^{\frac{\mu}{\mu+1}}. \end{aligned}$$

Thus, Algorithm 2 yields an order optimal reconstruction scheme for all  $0 < \mu \leq \mu_0$ . It follows again from [14, Theorem 2.1] that Algorithm 2 is also a regularization method for (1.1).  $\square$

## 5 Examples and numerical tests

In this final section, we present a simple test equation, in which the preliminary assumptions of Theorem 4.1 and 4.3 are satisfied. With help of this example we test the convergence order and the performance of the introduced adaptive algorithms. As a simple example of a linear problem (1.1) in which the operator  $A$  is in the class  $\mathcal{H}^2$ , we consider in  $X = L^2([0, 1])$  the Fredholm integral equation of the first kind (see [3, Example 12.4.1.]

$$Ax(t) = \int_0^1 k(s, t)x(s)ds = f(t), \quad 0 \leq t \leq 1, \quad (5.1)$$

with the kernel

$$k(s, t) = \begin{cases} t(s-1) & 0 \leq t < s \leq 1, \\ s(t-1) & 0 \leq s \leq t \leq 1. \end{cases}$$

The self-adjoint operator  $A$  maps  $L^2([0, 1])$  into the Sobolev space  $W^{2,2}([0, 1])$ . In the setting of boundary value problems, the application of  $A$  to  $x$  corresponds to the solution of the boundary value problem  $f''(t) = x(t)$  with homogeneous boundary conditions  $f(0) = f(1) = 0$ . It is well-known that the functions

$$e_k(t) = \sqrt{2} \sin(\pi kt), \quad k \in \mathbb{N}, \quad t \in [0, 1],$$

form an orthonormal basis of eigenfunctions of the operator  $A$  with corresponding eigenvalues  $\lambda_k = -(\pi k)^{-2}$ . Therefore, we have  $\|A\| \leq \pi^{-2}$  and  $\|(I - P_m)A\| = \|A(I - P_m)\| \leq (\pi(m+1))^{-2}$ . This implies in particular  $\pi^2 A \in \mathcal{H}^2$ .

In (5.1), we consider the two different right hand sides  $f_1$  and  $f_2$  given by

$$\begin{aligned} f_1(t) &= t^3(1-t)^3, \\ f_2(t) &= \frac{t^3}{3} - \max\left(0, t - \frac{1}{2}\right)^2 - \frac{t}{12}. \end{aligned}$$

The corresponding exact solutions of the inverse problem (5.1) are given by

$$\begin{aligned} x_1^\dagger(t) &= 18t(5t^3 - 10t^2 + 6t - 1) \quad \text{with } x_1^\dagger \in M_{\mu,\rho} \text{ for } 0 < \mu < 1.25, \\ x_2^\dagger(t) &= 2t - \text{sign}(2t - 1) - 1 \quad \text{with } x_2^\dagger \in M_{\mu,\rho} \text{ for } 0 < \mu < 0.25. \end{aligned}$$

## 5.1 Test of Algorithm 1

We test first Algorithm 1 for the two right hand sides  $f_1$  and  $f_2$ . As semiiterative method we chose the  $\nu$ -method with  $\nu = 1.5$ . Since the qualification of this method is  $\mu_0 = 3$  it can be applied in Algorithm 1 for both test examples. In this way we have  $\kappa_0 = 1$  and  $\kappa_2 = 6$ . Choosing  $\gamma = \frac{1}{2}$ , we can take  $\tau = 1.01 + \sqrt{\frac{13}{8}}$  as parameter for the discrepancy principle. We set  $\rho = 1$  and generate perturbed right hand sides  $f_{i,\delta}, i = 1, 2$  for different values of  $\delta > 0$ . Now, we use Algorithm 1 to compute the approximate solutions  $x_{i,n,K}^\delta$  of  $x_i^\dagger$ . The resulting errors and stopping indices of Algorithm 1 are displayed in Table 1, 2 and Figure 2, 3

Table 1: Results of Algorithm 1 to solve the test problem  $Ax_1 = f_1$ .

$\nu$	$\delta$	$n$	$K_n$	$K$	$\frac{\ x_{1,n,K}^\delta - x_1^\dagger\ }{\ x_1^\dagger\ }$	$\delta^{\frac{1.25}{1+1.25}}$
1.5	0.062500	6	37	12	0.49975111	0.21431100
1.5	0.031250	6	19	17	0.29238913	0.14581613
1.5	0.015625	7	125	20	0.21650878	0.09921257
1.5	0.007812	7	63	24	0.17715080	0.06750373
1.5	0.003906	8	435	45	0.10086226	0.04592920
1.5	0.001953	8	218	57	0.07100275	0.03125000
1.5	0.000977	8	109	80	0.04971398	0.02126234
1.5	0.000488	9	774	108	0.03362040	0.01446679
1.5	0.000244	9	387	147	0.02322422	0.00984313
1.5	0.000122	10	2784	203	0.01549616	0.00669722

Figure 2: The error and the stopping index  $K$  of Algorithm 1 compared to the expected error  $\delta^{\frac{1.25}{1+1.25}}$  and the expected stopping index  $\delta^{-\frac{1}{1+1.25}}$  of the first test problem  $Ax_1 = f_1$ .

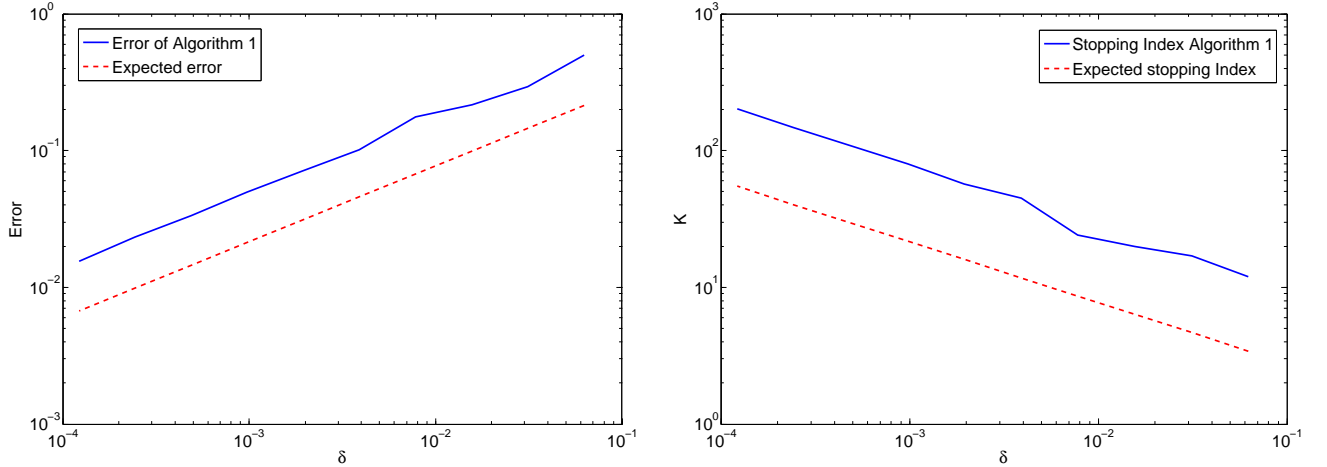
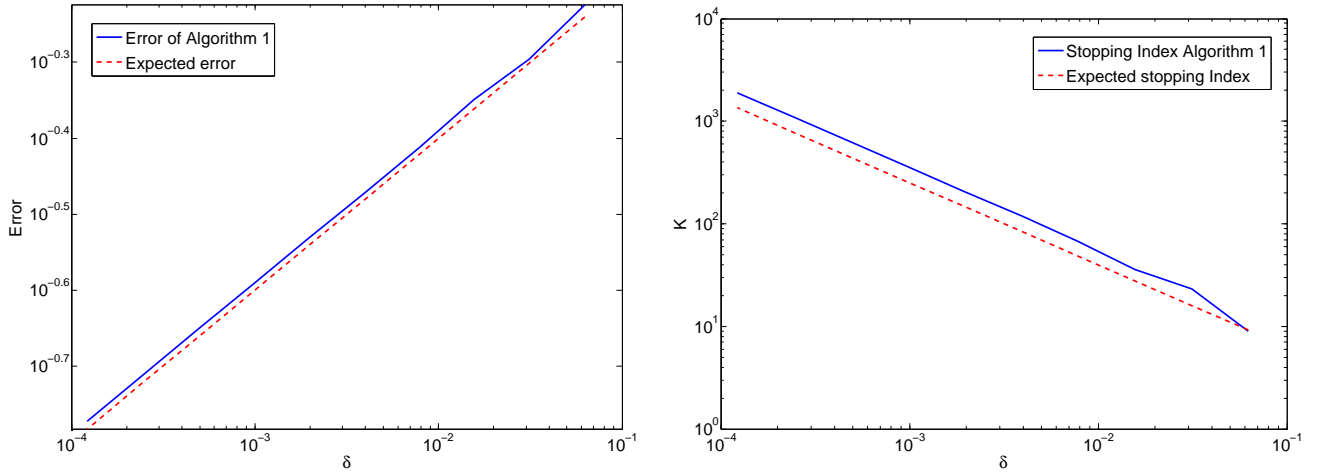


Table 2: Results of Algorithm 1 to solve the test problem  $Ax_2 = f_2$ .

$\nu$	$\delta$	$n$	$K_n$	$K$	$\frac{\ x_{2,n,K}^\delta - x_2^\dagger\ }{\ x_2^\dagger\ }$	$\delta^{\frac{0.25}{1+0.25}}$
1.5	0.062500	6	16	9	0.59696031	0.57434918
1.5	0.031250	7	105	23	0.50523819	0.50000000
1.5	0.015625	7	53	36	0.44800149	0.43527528
1.5	0.007812	8	366	68	0.38638037	0.37892914
1.5	0.003906	8	183	120	0.33629235	0.32987698
1.5	0.001953	9	1300	207	0.29364826	0.28717459
1.5	0.000977	9	650	361	0.25566471	0.25000000
1.5	0.000488	10	4678	625	0.22295988	0.21763764
1.5	0.000244	10	2339	1091	0.19402742	0.18946457
1.5	0.000122	11	17010	1901	0.16890368	0.16493849

Figure 3: The error and the stopping index  $K$  of Algorithm 1 compared to the expected error  $\delta^{\frac{0.25}{1+0.25}}$  and the expected stopping index  $\delta^{-\frac{1}{1+0.25}}$  of the second test problem  $Ax_2 = f_2$ .



## 5.2 Test of Algorithm 2

Now we test Algorithm 2 for the two right hand sides  $f_1$  and  $f_2$ . As semiiterative method we choose again the  $\nu$ -method with  $\nu = 1.5$ . As further control parameter, we choose  $\gamma = \frac{1}{2}$ . Also for Algorithm 2, we set  $\rho = 1$  and generate perturbed right hand sides  $f_{i,\delta}$ ,  $i = 1, 2$ , for different values of  $\delta > 0$ . With Algorithm 2 we then compute approximate solutions  $x_{i,n,K}^\delta$  of  $x_i^\dagger$ . The resulting errors for the two test problems are displayed in Table 3, 4 and Figure 4, 5.



Table 3: Results of Algorithm 2 to solve the test problem  $Ax_1 = f_1$ .

$\nu$	$\delta$	$n$	$K_n$	$K$	$\frac{\ x_{1,n,K}^\delta - x_1^\dagger\ }{\ x_1^\dagger\ }$	$\delta^{\frac{1.25}{1+1.25}}$
1.5	0.062500	6	37	8	0.68979661	0.21431100
1.5	0.031250	6	19	15	0.36601474	0.14581613
1.5	0.015625	7	125	19	0.23679445	0.09921257
1.5	0.007812	7	63	22	0.18993287	0.06750373
1.5	0.003906	8	435	33	0.14615533	0.04592920
1.5	0.001953	8	218	48	0.09181469	0.03125000
1.5	0.000977	8	109	59	0.06784866	0.02126234
1.5	0.000488	9	774	88	0.04481807	0.01446679
1.5	0.000244	9	387	114	0.03125762	0.00984313
1.5	0.000122	9	194	158	0.02144644	0.00669722

Figure 4: The error and the stopping index  $K$  of Algorithm 2 compared to the expected error  $\delta^{\frac{1.25}{1+1.25}}$  and the expected stopping index  $\delta^{-\frac{1}{1+1.25}}$  of the first test problem  $Ax_1 = f_1$ .

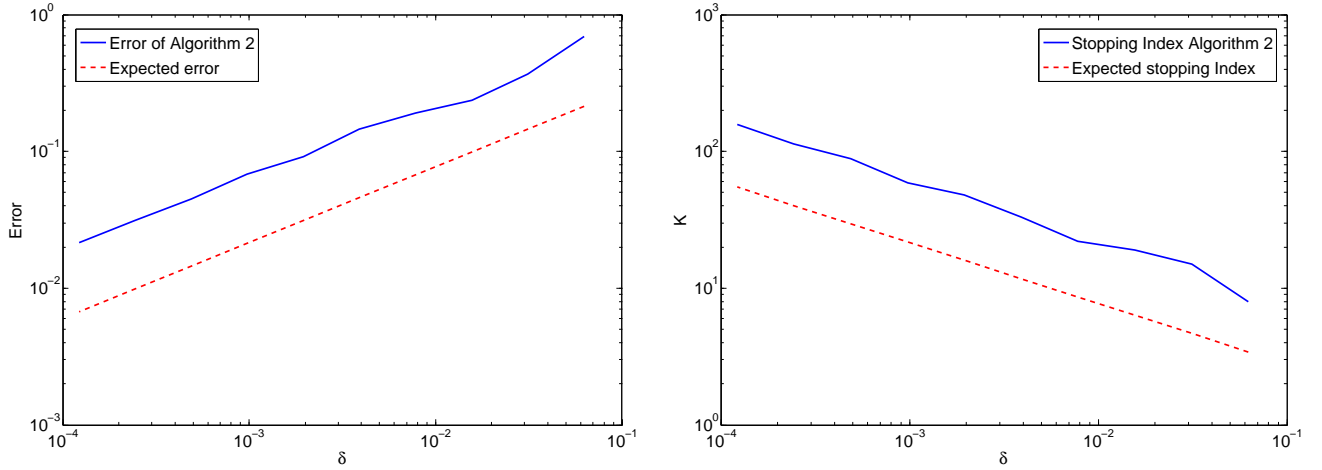
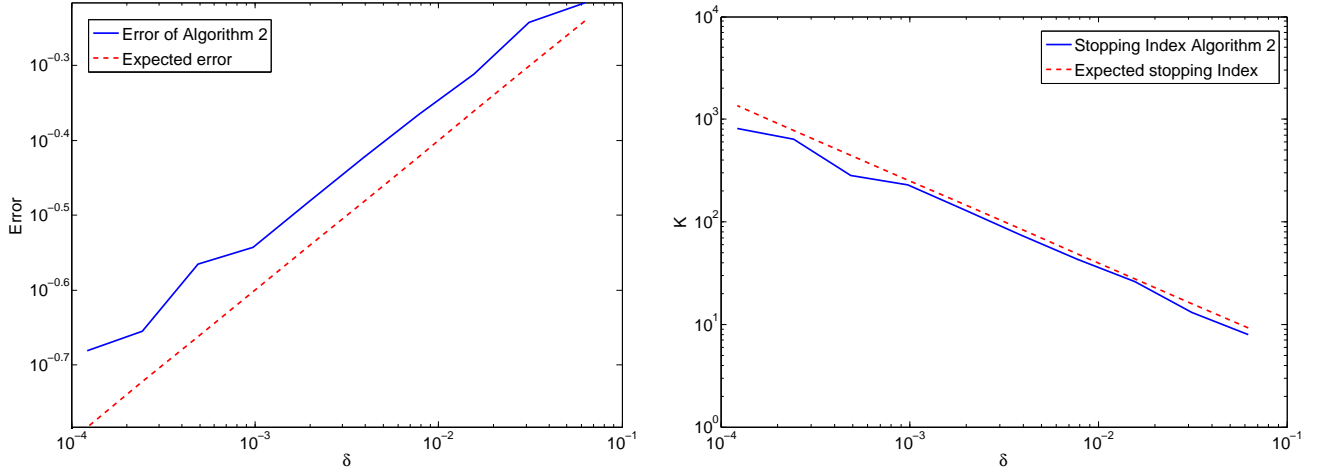


Table 4: Results of Algorithm 2 to solve the test problem  $Ax_2 = f_2$ .

$\nu$	$\delta$	$n$	$K_n$	$K$	$\frac{\ x_{2,n,K}^\delta - x_2^\dagger\ }{\ x_2^\dagger\ }$	$\delta^{\frac{0.25}{1+0.25}}$
1.5	0.062500	6	16	8	0.60790728	0.57434918
1.5	0.031250	7	105	13	0.57256321	0.50000000
1.5	0.015625	7	53	26	0.48809868	0.43527528
1.5	0.007812	8	366	43	0.43126730	0.37892914
1.5	0.003906	8	183	74	0.37805221	0.32987698
1.5	0.001953	9	1300	131	0.32892273	0.28717459
1.5	0.000977	9	650	228	0.28663978	0.25000000
1.5	0.000488	9	325	281	0.27204892	0.21763764
1.5	0.000244	10	2339	643	0.22139020	0.18946457
1.5	0.000122	10	1170	818	0.20848631	0.16493849

Figure 5: The error and the stopping index  $K$  of Algorithm 2 compared to the expected error  $\delta^{\frac{0.25}{1+0.25}}$  and the expected stopping index  $\delta^{-\frac{1}{1+0.25}}$  of the first test problem  $Ax_2 = f_2$ .



The graphs in Table 1 - 4 confirm the theoretical results on the order optimality of Algorithm 1 and 2. The parameters  $\tau$  and  $\gamma$  in the two algorithms are chosen conservatively such that the assumptions in Theorem 4.1 and 4.3 are satisfied. Neglecting these theoretical preconditions on the parameters it is possible to further improve the numerical results. In particular for Algorithm 1, a smaller choice of  $\tau$  yields better results for the approximation error  $\frac{\|x_{n,K}^\delta - x^\dagger\|}{\|x^\dagger\|}$ . Also, increasing the parameter  $\gamma$  leads to smaller discretization levels  $n$ , as compared to the ones displayed in Table 1 - 4, and reduces the numerical costs of the algorithms.

## References

- [1] *H. Brakhage*, On ill-posed problems and the method of conjugate gradients, In *Inverse and ill-posed problems*, Alpine-U.S. Semin. St. Wolfgang/Austria 1986, H. W. Engl and C. W. Groetsch, Eds., Notes Rep. Math. Sci. Eng. 4 (1987), pp. 165–175.
- [2] *T. S. Chihara*, An Introduction to Orthogonal Polynomials, Gordon and Breach, Science Publishers, New York, 1978.
- [3] *L. M. Delves, J. L. Mohamed* Computational Methods for Integral Equations. Cambridge University Press, 1985.
- [4] *H. W. Engl, M. Hanke and A. Neubauer*, Regularization of inverse problems, Kluwer Academic Publishers, Dordrecht, 1996.
- [5] *W. Erb*, Accelerated Landweber methods based on co-dilated orthogonal polynomials, Numer. Alg. (2014), DOI: 10.1007/s11075-014-9842-z,
- [6] *M. Hanke*, Accelerated Landweber iterations for the solution of ill-posed equations, Numer. Math. 60 (1) (1991), pp. 341–373
- [7] *A. K. Louis*, Inverse und schlecht gestellte Probleme, Teubner-Verlag, Stuttgart, 1989.
- [8] *P. Maass, S. V. Pereversev, R. Ramlau, S. G. Solodky*, An adaptive discretization for Tikhonov-Phillips regularization with a posteriori parameter selection, Numer. Math. 87 (2001), pp. 485–502.

- [9] V. A. Morozov, On the solution of functional equations by the method of regularization, Soviet Math. Doklady 7, (1966), pp. 414–417.
- [10] A. S. Nemirovskii, B. T. Polyak, Iterative methods for solving linear ill-posed problems under precise information. II. Engrg. Cybernetics 22 (4) (1984), pp. 50–56.
- [11] S. V. Pereverzev, Optimization of Projection Methods for Solving Ill-Posed Problems, Computing 55 (1995), pp. 113–124.
- [12] S. V. Pereverzev, E. Schock, On the adaptive selection of the parameter in regularization of ill-posed problems, SIAM J. Numer. Anal. 43 (5) (2005), pp. 2060–2076.
- [13] S. V. Pereverzev, S. G. Solodky, An Efficient Discretization for Solving Ill-Posed Problems, Lect. Appl. Math. 32 (1996), pp. 643–649.
- [14] R. Plato, Optimal algorithms for linear ill-posed problems yield regularization methods, Numer. Funct. Anal. Optimization 11 (1990), pp. 111–118.
- [15] R. Plato, G. Vainikko, On the Regularization of Projection Methods for solving Ill-posed Problems, Numer. Math. 57 (1990), pp. 63–79.
- [16] A. Rieder, Keine Probleme mit inversen Problemen, Vieweg Verlag, Wiesbaden, 2003.
- [17] S. G. Solodky, A generalized projection scheme for solving ill-posed problems. *J. Inverse Ill-Posed Probl.* 7 (2) (1999), 185–200.
- [18] S. G. Solodky, E. V. Lebedeva, Bounds of information expenses in constructing projection methods for solving ill-posed problems, *Comp. Method Appl. Math.* 6 (1) (2006), 87–93.
- [19] S. G. Solodky, E. A. Volonets, Adaptive scheme of discretization for one semiiterative method in solving ill-posed problems, Ukrainian Mathematical Bulletin 7 (4) (2010), pp. 553–569.
- [20] A. N. Tikhonov, V. A. Arsenin, Solutions of ill-posed problems, Wiley, New York, 1977.